A Model for Knowledge-Based Pronoun Resolution

Stefan L. Frank (stefra@mpi.nl)
Discourse Studies, Tilburg University, and Center for Language Studies, University of Nijmegen
P.O. Box 310, 6500 AH Nijmegen, the Netherlands

Mathieu Koppen (koppen@nici.kun.nl)
NICI, University of Nijmegen

Leo G.M. Noordman (noordman@uvt.nl)
Discourse Studies, Tilburg University

Wietske Vonk (vonk@mpi.nl)
Max Planck Institute for Psycholinguistics and Center for Language Studies, University of Nijmegen

Several sources of information are used in choosing the intended referent of an ambiguous pronoun. The two sources considered in this paper are foregrounding and context. The first refers to the accessibility of discourse entities. An entity that is foregrounded is more likely to become the pronoun's referent than an entity that is not. Context information affects pronoun resolution when world knowledge is needed to find the referent.

The model presented here simulates how world knowledge invoked by context, together with foregrounding, influences pronoun resolution. It was developed as an extension to the Distributed Situation Space (DSS) model of knowledge-based inferencing in story comprehension (Frank, Koppen, Noordman, & Vonk, 2003), which shall be introduced first.

The Distributed Situation Space

A problem for computational models of story comprehension is the large amount of world knowledge required. This is dealt with by decreasing the size of the world in which stories take place. In the resulting microworld, two story characters exist, named Bob and Joe. Their possible activities and states are described using 14 propositions, such as *Bob is tired*, and *Joe wins*. Any microworld situation is a combination of (some of) the 14 propositions. Some situations are more likely than others, because propositions are not unrelated but put constraints on each other. There also exist constraints between consecutive situations, which are assumed to follow one another in discrete time steps. Using the within-situation and between-situation constraints, a *microworld description* of 250 consecutive example situations was constructed, from which the required world knowledge was extracted.

The DSS model uses a distributed representation for propositions, which is obtained by training a Self-Organizing Map (Kohonen, 1995) on the microworld description. After training, each proposition p is represented by a vector $\mu(p) \in [0,1]^{150}$ in a 150-dimensional space, called the *situation space* because vectors in the space represent microworld situations.

The advantages of this representation are threefold. First, propositions can be combined using Boolean operators, so any microworld situation vector can be computed. Second, the unconditional probability that a proposition p is the case, denoted $\tau(p)$, can be computed from its representation. Third, the probability that a proposition p occurs, given that q occurs at the same moment, can be computed from $\mu(p)$ and $\mu(q)$. This value, denoted $\tau(p|q)$, is called the *belief value* of p given q, since it indicates to what extent proposition p is believed if q is known to be the case. Using Markov random field theory, it is possible to compute the belief value $\tau(p_{t\pm1}|X)$ that p occurs directly preceding or following some *context situation* denoted by X. This computation uses knowledge about the constraints between *consecutive* situations, which is also based on the microworld description.

The Pronoun Resolution Model

The DSS model simulates knowledge-based inference without using textual cues such as pronouns. Here, it is extended with a pronoun resolution process. We shall only consider texts containing *he wins*, with two possible referents (Bob and Joe) for 'he'. There can also be a context statement X. If X stands for *Bob is tired and Joe is not* and precedes *he wins*, the text reads *Bob is tired and Joe is not, so he wins*.

Without context, the belief value of proposition p is $\tau(p)$. If a context X is given, this value changes to $\tau(p_{t\pm1}|X)$. The influence that X has on p is the difference between the two: $\tau(p_{t\pm1}|X) - \tau(p)$. The *context preference* for p is the difference between its effects on p and on q: $[\tau(p_{t\pm1}|X) - \tau(p)] - [\tau(q_{t\pm1}|X) - \tau(q)]$. The *context strength* is the absolute value of the context preference. The stronger the context, the more it prefers either Bob or Joe as the pronoun's referent.

In the DSS model, context information changes a situation vector, resulting in an inference. In the current model, pronoun resolution is viewed as a special case of this, where the possible inferences are restricted. In our simulations, the possibilities are *Bob wins* and *Joe wins*, so a choice has to be made between these two propositions, represented by points $\mu(B)$ and $\mu(J)$ in situation space. If vector Y represents the current interpretation of *he wins*, the pronoun resolution process moves Y towards either $\mu(B)$ or $\mu(J)$. This is done by letting both $\mu(B)$ and $\mu(J)$ act as 'attractor points' and 'pull' at Y.

This process is defined by a pair of differential equations. When the initial position of vector Y and its initial velocity are given, the equations can be solved, resulting in a path through situation space that Y follows over processing time. This path ends when Y reaches either $\mu(B)$ or $\mu(J)$. Processing time of the model is defined as the time over which the differential equations are evaluated and is expressed in abstract 'units'.

Whether Y ends up at $\mu(B)$ or at $\mu(J)$ depends both on foregrounding and on context. Foregrounding is incorporated into the initial position of Y, and context information determines its initial velocity. The precise effect of context follows directly from the DSS model.

Both during and after the pronoun resolution process, the belief values $\tau(B|Y)$ and $\tau(J|Y)$ indicate the extent to which the pronoun is instantiated to Bob and Joe respectively. The entity corresponding to the highest belief value is considered to be the chosen referent.

Results

When a pronoun is read it is partially instantiated to the foregrounded entity, before context information comes into play (Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000). Figure 1 shows that the model simulates this when processing the sentence *Bob is tired and Joe is not, so he wins,* where Bob is foregrounded because he is mentioned first. At the onset of processing time, the pronoun has already been partially instantiated to Bob. As processing continues, context information exerts enough influence to change this into 'Joe wins'.

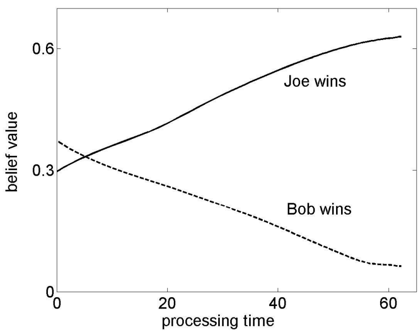

Figure 1: Belief values of *Bob wins* and *Joe wins* during processing of *Bob is tired and Joe is not, so he wins*.

All results below are obtained from processing the statement *he wins* in 56 different contexts and with three different foreground conditions (Bob, Joe, neither). Context strengths ranged from .0003 to .07. After each simulation, the belief value of either *Bob wins* or *Joe wins* was clearly the largest, so a referent was selected. The foregrounded entity was more likely to be chosen: If Bob was foregrounded, he was chosen 43 times out of 56; without foregrounding he was chosen 28 times; with Joe foregrounded Bob was chosen 13 times. Context preference also has an effect. Without foregrounding, the chosen referent followed the context in 54 items.

Leonard, Waters, and Caplan (1997) had participants read sentences with an ambiguous pronoun. A sentence was *congruent* if foregrounding and context pointed towards the same entity, and *incongruent* otherwise. It was found that (1) the correct entity (consistent with context) was chosen more often in congruent sentences than in incongruent ones; (2) the correct entity was chosen more often with stronger context; (3) reading times were shorter for congruent sentences than for incongruent ones; (4) reading times were shorter with stronger context. The model predicts all four of these effects: (1) no errors are made with congruent items, but 30 (out of 56) times with incongruent items; (2) correctly resolved pronouns had an average context strength of .042, compared to .012 for the others; (3) the average processing time for congruent items was 49.7, compared to 76.8 for the others; (4) without foregrounding, the correlation between context strength and processing time was –.81.

Conclusion

The model shows how the interplay between foregrounding and context information determines the referent of a pronoun. It is completely compatible with the DSS model of knowledge-based inferences, showing how knowledge-based pronoun resolution can be regarded as a special case of general knowledge-based inference.

References

Arnold, J.E., Eisenband, J.G., Brown-Schmidt, S., & Trueswell, J.C. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eyetracking. *Cognition*, *76*, B13-B26.

Frank, S.L., Koppen, M., Noordman, L.G.M., & Vonk, W. (2003). *Modeling knowledge-based inferences in story comprehension*. Manuscript submitted for publication.

Kohonen, T. (1995). *Self-Organizing Maps*. Berlin: Springer.

Leonard, C.L., Waters, G.S., & Caplan, D. (1997). The influence of contextual information on the resolution of ambiguous pronouns by younger and older adults. *Applied Psycholinguistics*, *18*, 293-317.